Discrimination in the Dutch Labour Market: A Correspondence Test of Race and Education

Dennis van Doornik - 625516

Erasmus Universiteit Rotterdam

Master Thesis - FSWS575

Wordcount Thesis: 9671

Wordcount Appendices: 1970

Dr Bonnie French & Dr Jennifer Holland

June 19, 2022

Table of Contents

List of Tables

List of Figures

Acknowledgements

Foremost, I wish to express my appreciation to my thesis supervisor, Dr Bonnie French, who guided me throughout the research process. Her expertise in inequality and passion has enabled me to conduct this study. I also owe a debt of gratitude to Tarik Bel Hadi, who advised me on the structure of the study and provided insights into the Dutch labour market. Lastly, I would like to thank my thesis group, which provided support and periodic feedback.

Abstract

This study examines racial discrimination in the Dutch labour market and the influence of education. In a correspondence test, this study sent fictitious paired applications of black and white male candidates to entry-level jobs, distinguishing between semi-skilled jobs and jobs requiring a college education. The resumes were assigned a photo of a black or white man whose skin colour served as a proxy to manipulate race. Three hundred applications were sent to 150 job vacancies. This study found significant discrimination against black applicants, who must submit 28 percent more job applications than equally qualified white applicants to receive a callback from employers. These racial disparities demonstrate the significance of skin colour in the Dutch labour market. So, the results contradict the ideology of a just world in which every person has equal opportunities. Moreover, the data did not support that the difference in callbacks between white and black candidates is smaller when applying for jobs requiring a college education than for semi-skilled jobs. So, this study did not find that a higher-level educational degree significantly decreased racial discrimination against blacks. As such, skin colour and race provide a crucial area for future research in the Dutch context. I discuss the implications for theory and practice.

Keywords: Correspondence test, Discrimination, Labour market, Race, Skin colour

Introduction

According to the meritocratic ideology, differences between people in terms of their social position, income, and social network are justified as long as everyone, regardless of their origin, has the same opportunities to develop their individual capabilities (de Beer & van Pinxteren, 2016). As in many Western countries, this ideology is widely embraced in the Netherlands (Steijn et al., 2016). However, there are still many inequalities among various groups in Dutch society (de Beer & van Pinxteren, 2016). A popular notion is that education is the great equaliser. From an early age, children are taught that education is the door to success. Statistics support this idea; unemployment rates decrease and wages increase with increasing levels of education (Centraal Bureau voor de Statistiek, 2021; Centraal Bureau voor de Statistiek, 2020a). However, these statistics do not account for additional factors that might influence a person's job opportunities, such as racism or discrimination. These factors are central to this study. Specifically, I will take a closer look at racial discrimination in the Dutch labour market and the influence of education.

One of the earliest studies on racism in the Netherlands is from pioneer Essed (1991). She gave several black citizens a voice; they generally agreed that racism is inherent within the social system. Even though this finding cannot be generalised, Essed argues that it calls for more research into racism within the Netherlands. However, 30 years later, extensive research in this area is still absent. More recently, Hondius (2014) conducted numerous interviews with black Dutch men and women to examine their experiences in the Netherlands. Many interviewees in her study were rejected during job applications and suspected that their skin colour played an important role in their rejection. The current study responds to the call of Essed (1991) and the experiences of the respondents in the study of Hondius (2014) by examining if employers favour white applicants when similar black and white applicants apply for the same job.

The Dutch case is particularly interesting because the Netherlands seems to regard itself as inherently non-racist; terms like race and skin colour are generally taboo (Hondius, 2014; Wekker, 2016). Some argue that differential treatment by race is something of the past and that racial discrimination has been eliminated by anti-discrimination legislation, public pressure, and diversity management strategies (Sabharwal, 2014; Veenman, 2010). However, as mentioned above, several qualitative studies report that systematic racial disadvantage is to be expected in all areas of Dutch society, including recruitment processes due to employer prejudice (Essed, 1991; Essed & Hoving, 2014; Hondius, 2014; Wekker, 2016; Wekker & Lutz, 2001).

Race is generally associated with phenotypic characteristics of which skin colour is deemed the most prominent feature (Omi & Winant, 2014; Sesardic, 2010). Therefore, it is important to gain a better understanding of skin colour and the potential implications on job opportunities. However, current data limitations make it difficult to empirically test whether employers find people's phenotypic features, such as skin colour, either relevant or irrelevant. Here I extend current work by explicitly examining labour market discrimination based on skin colour to assess to what extent this physical characteristic affects job opportunities in the Netherlands.

Field experiments in the United States consistently show that white applicants receive more callbacks from potential employers than black applicants after submitting a job application (Baert, 2018; Bertrand & Duflo, 2017; Quillian et al., 2017; Riach & Rich, 2002). These findings are consistent with the earlier mentioned experiences of black Dutch citizens in the study of Hondius (2014). Therefore, I hypothesise that black applicants receive fewer callbacks than white applicants within the Dutch labour market.

Furthermore, I hypothesise that education does not serve as the great equaliser for racial discrimination in the Dutch labour market. Bovenkerk et al. (1995) found that there is

less discrimination for jobs that require a higher level of education in the Netherlands. As a result, they argued that education is essential in reducing discrimination and enhancing the position of minorities in the labour market. Nevertheless, there was still a considerable amount of discrimination in the applications for jobs that require a college education. Therefore, I hypothesise that employers will discriminate less in terms of callbacks between white and black candidates for jobs requiring a college education than for semi-skilled jobs. However, black applicants who apply for jobs requiring a college education will still receive fewer callbacks than similar white applicants.

Further, because of the limited timeframe of this research and to ensure sufficient statistical power, I determined that I had to choose whether to conduct this study for either men or women. In the end, I decided to study men because I identify as one myself and because this group is quantitatively the most involved in the Dutch labour market[1]. Accordingly, this study seeks to answer the following research question: "*To what extent does skin colour influence the likelihood of receiving a callback from potential employers of black and white male applicants within the Netherlands? And what is the impact of different levels of education on this effect*?".

Field experiments are widely considered the most appropriate way to identify racism in recruitment (Bertrand & Duflo, 2017; Riach & Rich, 2002). This approach provides information about natural hiring procedures, and it allows to isolate the effects of applicants through matching and random assignment (Pager, 2007; Pager & Shepherd, 2008). Therefore, I employed a field experiment methodology by conducting a correspondence test. By holding constant levels of human capital (i.e., work experience and educational attainment) while varying the signals of skin colour via different photographs on the resume, I can demonstrate

[1] This decision is explained in more detail in the theoretical framework. The discussion also highlights the importance of conducting similar research for women and other gender identities.

how these signals affect racial discrimination during job applications.

To the best of my knowledge, I am the first to examine racial discrimination based on skin colour in the Dutch labour market through photographs. By exploring the influence of skin colour on job opportunities for black and white men, I attempt to improve the understanding of racial discrimination within the Dutch labour market. Another contribution of this research is examining the extent to which levels of racism vary between different educational levels. I seek to advance the awareness of the possibilities and restrictions of education in reducing inequality. These insights potentially give policymakers and employers new perceptions about a significant yet often overlooked group of the Dutch population.

Theoretical Framework

The "just world hypothesis"

An important grounding of the ideology of equality in the Netherlands can be described with the "just world hypothesis" (Lerner, 1980; van den Broek, 2014). This theoretical presumption is based on the ideology that individuals get and achieve what they deserve. This idea validates and results in the concept of equality, suggesting that every person has equal opportunities. However, the just world hypothesis denies the reality of inequality.

The just world hypothesis treats the status quo as just, equal, fair, and inevitable (Lerner, 1980; Jost & Hunyady, 2003; van den Broek, 2014). The notion that the world is as it is meant to be provides people with existential security (van den Broek, 2014). This could explain where the ideology of the just world comes from. Individuals have a need to feel comfortable, even if it means ignoring or justifying inequality (van den Broek, 2014).

The ideology of the just world hypothesis is not unique to the Netherlands; it is a philosophy that is deployed worldwide (Duru-Bellat & Tenret, 2012). Yet, societies in certain countries believe more in the just world hypothesis than others. Researchers did a cross-

country analysis among dozens of countries and asked citizens to what extent they thought people got what they deserved (Duru-Bellat & Tenret, 2012; Weinberg et al., 2021). These studies show a relationship between GDP and the belief in a just world. Countries with a higher national income are generally more supportive of the just world hypothesis. Moreover, results show that income inequality generally parallels with meritocratic beliefs (Duru-Bellat & Tenret, 2012). Societies seem to believe less in a just world in countries with more income inequality. Thus, a country's financial position likely influences the perception of fairness in the social system. These findings are fairly self-evident; societies with relatively high income and little income inequality can more easily claim that the world is fair and just.

Compared to other European countries, the Netherlands scores above average in terms of the beliefs in a just world (Weinberg et al., 2021). These convictions also seem to be reflected in the Dutch debate about race and ethnicity.

Ethnicity and race in the Dutch context

The term race is utilised in several ways in research and is therefore susceptible to interpretation (Essed & Hoving, 2014). Because there is no consensus about the definition, I will first elaborate on the concept of race.

Race is frequently associated with biologically based human characteristics, so-called "phenotypes", which concern physical attributes such as skin colour, hair type, and facial features (Omi & Winant, 2014; Sesardic, 2010). These visual differences have been used for centuries to classify people into racial groups (Adelman, 2014). There is the continuing idea of thinking that race is fixed or objective. Yet, people disregard that the selection of human features for racial distinctions is a social and historical process (Omi & Winant, 2014; Wekker, 2016). So, race is a social construct; each person learns a form or version of racial classification, generally without explicit and conscious teachings (Omi & Winant, 2014)[2].

[2] Due to the objective of this study, I will generally use a simplified conception of race, which refers to physical or phenotypical characteristics such as skin colour.

Race is “real” in the sense that it has real consequences for a person’s perception of themselves and their opportunities in life (Frankenberg, 1993; Wekker, 2016). People form associated attributes with a racial group, which, often unconsciously, can lead to discrimination against others (Goffman, 1963; Kang et al., 2016).

In Dutch science and public policy, alternatives to racial terms have ruled the discourse for the past decades (Hondius, 2014). The term race has not been commonly employed in the Netherlands since World War II (Nimako & Willemsen, 1993; Wekker, 2016). Instead, terms such as ethnicity, minority, and the dichotomy of Western versus non-Western immigrant have been applied in Dutch academia and politics (Hondius, 2014). These presumably innocent definitions are racialising people without the need to express “distasteful” racial wording (Wekker, 2016; Wekker & Lutz, 2001).

An infamous example is the use of the terms autochtoon versus allochtoon. Autochtoon means people from the Netherlands, and allochtoon refers to those who come from elsewhere (Wekker, 2016). However, within the group of autochtonen, there are many with foreign ancestors who, because of their white skin colour, successfully manage to claim that they are Dutch (Essed, 1991; Wekker, 2016). In contrast, allochtonen are the people who cannot make this claim because of their skin colour (Essed, 1991; Wekker, 2016). “Non-natives of colour” remain allochtonen, never fully being included as being Dutch. Thus, being an autochtoon can be seen as being white. Another example is found within the concept of ethnicity. Ethnicity signifies the social system of differences between individuals based on country of origin, language, religious beliefs, and cultural background (Okazaki & Sue, 2016). However, in the Netherlands, a distinction is often made between Western and non-Western ethnicities, usually meaning white versus black (Hondius, 2014; Wekker, 2016).

In other words, the Dutch seem to avoid mentioning race or skin colour by creating and using alternative terms. The Netherlands appears to act like a country that is too tolerant

to talk about skin colour or, even worse, racial discrimination (Hondius, 2014). There is a more or less silent agreement not to discuss potential racial problems (Hondius, 2014). Wekker (2016, p. 80) argues that the Dutch white sense of the self can be summarised with the following quote: "We are a small nation, innocent; we are inherently anti-racist". This sense of self preserves the idea that the Dutch are colour blind, making a term such as "race" taboo (Wekker, 2016). Wekker (2016) also argued that people felt guilt, anxiety, and fear when they were confronted with terms such as race, skin colour, and the specific naming of white or black instead of the usual terms like allochtoon/autochtoon or Western/non-Western immigrant. This exposes the tensions of using racial wording in a nation that prides itself on being anti-racist.

In other countries, the discussion of race and physical characteristics such as skin colour is less taboo[3]. The United States is a good example of this. The US builds on generations of knowledge about the struggle against racism in society (Essed, 1991). Race and ethnicity are also frequently grouped and used interchangeably as the same construct in discrimination and broader social science research in the US (Chrobot-Mason, 2004). Later in the report, I will elaborate on how the US conducts research based on skin colour and racism.

As mentioned earlier, it is uncommon for anyone to talk directly about a potential penalty based on skin colour in the Dutch discourse. An interpretation of the limited use of race in Dutch studies can be that people's physical features are deemed irrelevant and meaningless. The experiences and identities of people based on their skin colour are considered "unimportant", as are their experiences of racism (Hondius, 2014; Wekker, 2016). Moreover, racial issues might be regarded as an outdated problem without any relevance in the 21st century (Essed & Hoving, 2014). Some research even supports these ideas by arguing

[3] There are also contexts where it is prohibited to study race, such as France. As in the Netherlands, the discussion in France is centred around alternative terms such as immigrants, the poor, and the place of residence (Amiraux & Simon, 2006). Because this literature is generally colour-blind, I have not included it in this study.

that skin colour or race is not of great significance in the Netherlands (Veenman, 2010). However, ignoring race makes acknowledging and examining racism a significant challenge.

In this research, I will respond to this critical gap in the literature by exploring to what extent race matters in the Netherlands. More specifically, I will conduct a field experiment where I will strip away socio-cultural signifiers to examine whether the physical identifier of skin tone affects a person's opportunities[4]. Put differently, I intend to use skin colour as a proxy for race to assess whether it challenges the just world hypothesis in the Netherlands.

Racial discrimination in the Dutch labour market

Again and again, research in the Dutch context shows that the place of birth and social context play a significant factor in what is allotted to a person. Since the 1970s, a great body of research has emerged in the Netherlands on the existence of labour market discrimination based on ethnicity. These studies, using correspondence tests and in-person audits, show similar results; the discrimination of ethnic minorities during job applications (e.g., Andriessen et al., 2012; Blommaert et al., 2013; Bovenkerk, 1977; Bovenkerk et al., 1995; Thijssen et al., 2020). Dutch applicants were invited for a job interview up to 40% more often than applicants with a non-Western background (Moroccan, Turkish, Surinamese or Antillean). Nevertheless, these studies have minimal or no mention of skin colour nor racial discrimination.

Even though there is limited research on racial discrimination within the Netherlands, there are several qualitative studies that recognise race and the ways in which it shapes the lives of people. One of the first studies is by Essed (1991), who conducted interviews with black women in the Netherlands. These women pointed out that they expected racism in practically all interactions with whites. They also stated the difficulty of covert racism: "No one ever said to me: you cannot do that because you are black". Essed (1991) argues that the

[4] I describe the details of the procedures and implementation later in the methods section.

discourse of racism is changing and that it is increasingly infused with the idea that blacks have a cultural deficit. Similarly, van den Broek (2014) states that "old" forms of expressing overt hatred and violence aimed at black people are considered socially unacceptable nowadays. Nevertheless, more subtle forms of racial discrimination, such as exclusion and marginalisation, are barely recognised as racism. Everyday marginalisation of blacks is associated with supposed cultural differences, making racial exclusion seem just and fair (Essed & Hoving, 2014; van den Broek, 2012, 2014).

Moreover, Hondius (2014) interviewed black people in the Netherlands and asked questions about their skin colour. Almost all of the 72 interviewees spoke about racism in the labour market. They were certain that their skin colour was one of the determinant factors why they were not selected in previous job applications. They experienced that white people had negative expectations of them, such as coming late, being loud, dressing improperly, not mastering the Dutch language, and being unintelligent. However, similar to the study by Essed (1991), these experiences were never explicitly expressed by whites during job applications. To respond to these experiences, I will examine racial discrimination within the Dutch labour market. In particular, I will assess whether an applicant's skin colour influences their job opportunities.

Much of the research on race and racism has been conducted within America. As a result, a considerable amount of data is available regarding the representation of black people in American society. These statistics reveal racial gaps in unemployment; blacks are twice as likely to be unemployed than whites (Austin, 2013; Cancio et al., 1996; Williams & Wilson, 2019). These inequalities in the American labour market have changed little since 1980. Even though displays of explicit prejudice have been declining, the statistics of implicit prejudice and stereotypes appear to have changed barely (Bobo et al., 2012; Devine & Elliot, 1995; Dovidio et al., 2010). These observations are similar to the statements of Essed (1991) and

van den Broek (2014) that racism has taken more covert and subtle forms (Bonilla-Silva, 2016; Dovidio et al., 2010; Kinder & Sears, 1981).

A large and ever-growing number of field experiments investigating discrimination have been published in the last few decades. Most of these studies respond to job vacancies with fictitious resumes (Bertrand & Duflo, 2017). To manipulate the perception of race or ethnicity, they generally assign a certain name to these resumes. With a few exceptions, these experiments offer empirical evidence of discrimination against racial minorities during job applications - although there are substantial differences in the magnitude of the effects across studies (Baert, 2018; Bertrand & Duflo, 2017; Quillian et al., 2017; Riach & Rich, 2002). A meta-analysis of field experiments conducted in the United States shows that whites receive, on average, 36% more callbacks from employers than equally qualified blacks during job applications (Quillian et al., 2017). As mentioned earlier, these findings are consistent with the experiences of black people in the Netherlands reported in the study by Hondius (2014). Therefore, I will examine whether applicants with a black skin colour receive fewer interview requests than their otherwise equivalent white counterparts.

Hypothesis 1: Black applicants receive fewer callbacks than equally qualified white applicants during job applications within the Dutch labour market.

The intersection of race and education for job opportunities

Bertrand and Mullainathan (2004) demonstrated that race also impacts the benefit of a better resume. In their study, applicants with a better resume had completed a certification degree, more work experience, fewer gaps in their employment history, and received certain honour awards. They showed that white applicants with higher quality resumes got 30% more callbacks, whereas having a higher quality resume for black applicants had a much smaller

positive effect. While one might expect that better credentials would eliminate employers' prejudices that black candidates lack certain skills, their results do not demonstrate this.

Furthermore, a meta-analysis of European studies showed that candidates face significantly more discrimination in the hiring process for low-skilled jobs than for high-skilled ones (Flage, 2019). Similarly, Bovenkerk et al. (1995) conducted a study within the Netherlands and found that the higher the educational level of the applicants, the less discrimination they encounter. Specifically, they showed that there is less discrimination when applying for jobs requiring a college education than for semi-skilled level jobs. Yet, Bovenkerk and colleagues found that there was still a substantial level of discrimination against applicants with higher-level education. These findings contradict the earlier mentioned "just world hypothesis" and suggest that rather than serving as the great equaliser, a higher-level educational degree does not entirely eliminate employers' biases against blacks. Nevertheless, it still seems that an applicant's educational level significantly affects their chances of getting a job in the Netherlands. Therefore, I will adopt an intersectional approach and examine the intersection of race and education.

An intersectional approach is important because it can help to gain a better understanding of the complexity of identities and the potential patterns of oppression that people face within society (Gharib, 2022; Harris & Leonardo, 2018). Intersectionality looks at multiple interactions of social identities that mutually influence social opportunities to identify overlapping mechanisms of privilege and marginalisation (O'Donnell & Richardson, 2020). For example, if I were to examine employment opportunities only through the lens of race, I would not be able to recognise how other aspects, such as education, might influence racial discrimination. Thus, an intersectional approach assists in advancing the understanding of job opportunities within different contexts.

The previously mentioned studies provide an empirical motivation to investigate the

intersection of race and education, but these studies offer little to no theoretical justification. Therefore, I will give four theoretical explanations for potential differences in racial discrimination between educational levels.

Firstly, research suggests that employers focus more on candidates' skills for jobs that require a college education (e.g., account manager or finance and control employee) because these positions generally involve more responsibilities and require special skills (Flage, 2019). In contrast, for semi-skilled jobs (e.g., receptionist or sales assistant), employer criteria appear to be less focused on skills but more on physical appearance and personality-based attributes (Flage, 2019; Moss & Tilly, 2003; Pager et al., 2009). Because of this, features such as skin colour may be more decisive for semi-skilled jobs than jobs requiring a college education.

Secondly, the difference in discrimination between the education levels could be explained by the fact that employers are trading skin-tone status for educational status. Employers might perceive a college education as a signal to disconfirm the generalised negative perceptions about racial minorities, whereas a lower level of education might be regarded as a signal that is insufficient to disconfirm these negative perceptions (Biernat & Kobrynowicz, 1997; Kerckhoff et al., 2001).

Thirdly, the difference in hiring discrimination might be related to labour market tightness. Evidence suggests that there is less discrimination in the selection process of jobs that are difficult to fill (Baert et al., 2015). In other words, the smaller the supply of labour, the less selective employers can be, and the less likely characteristics that are not directly job-relevant, such as racial background, will play a role (Sociaal en Cultureel Planbureau, 2010). In the Netherlands, more people are qualified to work semi-skilled jobs than jobs requiring a college education (Ministerie van Onderwijs, Cultuur en Wetenschap, 2021). Because of this difference, employers might be more selective towards and discriminatory against candidates in the hiring process for semi-skilled jobs than jobs requiring a college education (Bovenkerk

et al., 1995).

The last explanation might be that employers at higher levels are more likely to recognise the value of diversity. Semi-skilled jobs involve more routine tasks where diversity generally does not provide direct benefits (van Knippenberg & Schippers, 2007). In contrast, jobs that require a college education will be more focused on knowledge-intensive tasks. It is precisely for these kinds of tasks and jobs that diversity can add value and improve business performance (Homan et al., 2007; van Knippenberg et al., 2004; van Knippenberg & Schippers, 2007).

Combining earlier arguments, I hypothesise that there is less racial discrimination against blacks for jobs requiring a college education than for semi-skilled level jobs. So, the callbacks during job applications among black and white candidates will be more similar when applying for jobs requiring a college education compared to semi-skilled ones. Nevertheless, the different educational requirements for these jobs will not eliminate discrimination in terms of callbacks between black and white applicants.

> *Hypothesis 2:* The difference in callbacks between white and black candidates will be smaller when applying for jobs requiring a college education than for semi-skilled jobs within the Dutch labour market. However, black applicants who apply for jobs requiring a college education will still receive fewer callbacks than similar white applicants.

Furthermore, intersectionality is an instrument that assists scholars in developing a more thoughtful and explicit understanding of why they used certain characteristics and intersections (Griffith, 2012; Harris & Leonardo, 2018). One of these intersections is that of gender. Scholars regard race and gender as the broad structural forces that fundamentally shape people's economic and social opportunities (Griffith, 2012; Laveist, 1996; Warner &

Brown, 2011). Gender is thus a social category that is critical to conduct a thorough and unbiased examination of the role of skin colour in people's labour market opportunities, as it potentially affects job chances via both social and institutional pathways.

Due to time constraints and to ensure sufficient statistical power to evaluate the effect of race, this study demanded a narrow and manageable focus. So, I was faced with the decision of whether to conduct this study specifically for men or women[5]. Even though I preferred to conduct this research for both men and women, I have ultimately opted to focus specifically on men for two reasons. Firstly, from my own positionality, it made sense to follow the intersectional tradition of examining and analysing men because I identify as one myself. Secondly, this study is the first to research the effect of skin colour on a person's job opportunities in the Netherlands. Therefore, I choose to study the impact on the group that is quantitatively the most involved in the labour market and possibly also the most affected[6]. However, I must state explicitly that the quantitative measurement is not an indicator of the importance and potential harm and burden for women and other gender identities. Thus, a critical note of this research is that the hypotheses are specifically focused on men. Later in the discussion, the importance of doing a similar study for other groups will be addressed in more depth.

Methods

To test the hypotheses, I have conducted a field experiment in the form of a correspondence test. Field experiments permit the examination of the causal effect of race on labour market outcomes in natural employment settings (Pager, 2007; Pager & Shepherd,

[5] To draw valid conclusions for both men and women or to exclude that gender had no effect would have required twice as much data. In addition, this would have made the data collection procedure even more challenging and time-consuming. However, this research was conducted as a thesis project with limited time to collect the data. Because of this restriction, I concluded that it was not possible to study both men and women.
[6] The most recent data of absolute labour force participation between men and women in the Netherlands is from 2020. Roughly 4,808,200 men participated in the labour market and approximately 4,233,300 women (Centraal Bureau voor de Statistiek, 2020b, 2020c). These figures refer to the active labour force between 15 and 75 years.

2008). In correspondence tests, two carefully matched, fictitious applicants apply to various job vacancies with cover letters and resumes to examine if one of the applicants is more likely to be invited for an interview than the other (Baert, 2018; Bertrand & Duflo, 2017; Riach & Rich, 2002).

In the Netherlands, it is customary to attach a picture to a job application (Di Stasio & Lancee, 2020; Lancee, 2019). I took advantage of this phenomenon and created two fictitious applicants, each with different photographs, one of a black man and one of a white man. These photos signalled the applicants' skin colour (and other phenotypic features), which functioned as a proxy for race. All features of the photos are kept as equivalent as possible (e.g., age, facial expression, hair length and colour, headpose, and background) to prevent other factors from influencing the results[7].

Other aspects of the resume were altered to similar characteristics so employers would not receive identical resumes. Both CVs had a Dutch-sounding name, ensuring that the origin of the candidates could not be deduced from the assigned name[8]. Consequently, differences in callbacks could be attributed to the physical features of the candidates. I also did not specify any other cultural markers (language, nationality, or religion) to isolate the effect of skin colour. Moreover, the applicants did not differ more than one year in age. Further, I kept the applicants' residence areas equivalent. Specifically, I used actual addresses but with non-existent house numbers. Both neighbourhoods were equivalent in terms of socioeconomic status to avoid discrimination based on class.

Moreover, because jobs are examined at two distinct levels, these fictitious black and white applicants differed in their level of education, creating four different resumes. This

[7] I used AI-generated photos to avoid detection and potential negative implications on actual people. All images generated by such a system are different, which made detection through the candidates' photos almost impossible. In addition, I purchased the rights for these pictures to prevent issues with personality rights or royalties.

[8] The first names were picked from the most common names of people born in the Netherlands between 1999 and 2004. The surnames were selected from the top 50 most common Dutch surnames.

study used two CVs for semi-skilled jobs and two for jobs requiring a college education[9] [10]. Applicants had post-secondary vocational education (level four) for semi-skilled level jobs, known as “middelbaar beroepsonderwijs - niveau vier (MBO)” in the Netherlands. Fictitious candidates who applied for a job requiring a college education possessed a vocational bachelor’s degree, referred to in the Netherlands as “hoger beroepsonderwijs (HBO)”.

To determine which vacancies to apply for, I consulted recruitment specialist Tarik Bel Hadi. Ultimately, based on the current situation of the labour market and the time and resource constraints of this study, we concluded to focus on the following occupations. For semi-skilled jobs, it was accessible to apply for jobs as financial administrator, receptionist, and sales assistant. In contrast, jobs requiring a college education were restricted to jobs in the positions of account management, finance, and human resources[11]. For a detailed explanation of these decisions, see Appendix A.

Furthermore, the candidates had just completed their studies and had no work experience. This background is beneficial because it suits many vacancies that do not demand a long employment history that might influence the employer’s perception (Campos-Vazquez & Gonzalez, 2020). A theoretical reason for examining this population is that the beginning of an individual’s career is a crucial period with potentially harmful long-term consequences for future occupations (Luijkx & Wolbers, 2009; Thijssen et al., 2020). This stresses the significance of obtaining a deeper understanding of the social obstacles at the beginning of a person’s career. However, virtually all vocational studies in the Netherlands involve one or more internships. Therefore, I included two internships that correspond to the program of the

[9] In terms of layout, the CVs of candidates applying for semi-skilled jobs differed from those applying for jobs requiring a college education. The CVs of candidates who applied for semi-skilled jobs were slightly less professional in terms of design. This distinction has also been used in similar studies to avoid detection (e.g., Andriessen, 2010, 2012).

[10] The resumes used in this study are not attached due to privacy considerations and potential implications for me as a researcher. However, to ensure replicability across context and time, I can be contacted to gain further insight into the specifics of the cover letters and resumes.

[11] To keep the resumes relevant to job positions, I adjusted the specialisation of the candidate's education to fit the vacancy in question.

respective study in the CVs.

During all applications, I submitted two matched resumes in response to each vacancy, both with either a post-secondary vocational education or vocational bachelor's degree, depending on the level of the job. Moreover, I attached a cover letter to each application[12]. For a comprehensive description of the procedures, see Appendix B.

The reactions of employers to the fictitious applicants were dichotomised. Invitations for an interview, questions for additional information, and callback requests were considered positive responses. Negative responses included non-responses and explicit rejections. Systematic differences in these responses to the fictitious candidates provided the measure of racism (Riach & Rich, 2002). Responses were captured via specific email addresses and telephone inboxes for each applicant[13].

For context considerations, I registered and coded the region where the job vacancy was located. I did this to examine whether racial discrimination differs between more or less urbanised areas. The distinction that I applied stems from the Dutch Federal Statistical Office, which is used to identify local differences in municipalities (Centraal Bureau voor de Statistiek, 2018). In particular, I made three distinctions between (a) jobs offered in the four largest municipalities (G4)[14], (b) vacancies in 40 other large and medium-sized municipalities (G32)[15], (c) and other relatively smaller municipalities (Thijssen et al., 2020).

Lastly, there were serious ethical concerns that applied to this study. The first issue is that participants were unaware that they were involved in an experiment. It was also impossible to attain informed consent, as informing participants would invalidate the study.

[12] For each occupation, two different motivation letters were created to be sent simultaneously with the CVs. Both the CVs and motivation letters were developed in consultation with recruitment specialist Bel Hadi.
[13] All calls were immediately sent to voicemail, which was the standard message of the mobile provider.
[14] The Hague, Utrecht, Rotterdam, and Amsterdam.
[15] Alkmaar, Almelo, Almere, Alphen aan den Rijn, Amersfoort, Apeldoorn, Arnhem, Assen, Breda, Delft, Deventer, Dordrecht, Ede, Eindhoven, Emmen, Enschede, Gouda, Groningen, Haarlem, Haarlemmermeer, Heerlen, Helmond, Hengelo, s-Hertogenbosch, Hilversum, Hoorn, Leeuwarden, Leiden, Lelystad, Maastricht, Nijmegen, Oss, Roosendaal, Sittard-Geleen, Schiedam, Tilburg, Venlo, Zaanstad, Zoetermeer, and Zwolle.

However, direct evidence of labour market discrimination is not obtainable by any other method (Bertrand & Duflo, 2017; McGinnity et al., 2009). Thus, it is almost impossible to measure how and under what circumstances skin colour influences employer behaviour without examining actual hiring decisions. Additionally, List (2009) argues that when a study benefits society and ensures anonymity and just treatment, the absence of informed consent is justifiable. To minimise the inconvenience for the employer or genuine applicants, I politely terminated the recruitment process when employers responded positively[16]. McGinnity et al. (2009) contended that it is not likely that any employer would endure substantial expenses or inconvenience with these procedures. Finally, the outcome of this study was treated confidentially (Riach & Rich, 2002)[17]. I have also attempted to keep the collection of names and contact details of employers to a bare minimum. I aimed to study labour market discrimination on a macro level, not discrimination by individuals on a micro level. Specifically, I focussed on collecting general characteristics to place clear definitional bounds on this macro-level space (e.g., the location, job position, and the required education of the vacancies).

In sum, I took all possible actions to minimise the individual risk and inconvenience for organisations and employers. The methodology and implemented procedures followed the best practices of numerous studies.

Results

In total, 300 applications were sent in pairs to 150 job advertisements. The data was collected over a period of six weeks. In these six weeks, the first two weeks were used to send out the applications and the remaining four weeks to collect responses from employers. After

[16] Specifically, when an employer contacted both applicants, I withdrew the application as promptly as possible (within one day). However, if they only reached out to one of the applicants, I waited two days to remove the application.

[17] No data related to any individual business was distributed to ensure that the results would not directly impact any participants. So, organisation identifying information was not made available to any third party at any stage, including my thesis group and supervisor.

this period, 14 pairs of matched job applications received no response and were counted as invalid. In 42 cases of the remaining 136 pairs of applications, neither of the candidates received a positive response. Thus, in 94 dyads, at least one of the candidates got a positive response. Table 1 shows a breakdown of the findings relating to the pairs of matched job applications.

Validity check

Firstly, I conducted a validity check to rule out that one of the resumes or letters got significantly more positive responses (Baert, 2018; Bertrand & Duflo, 2017; Bovenkerk et al., 1995). Moreover, I examined whether the time of sending the applications (first or second) affected employer responses. All three variables were not significant at the five percent level[18]. Therefore, this study found that unequal treatment was not due to the resumes, the cover letters, or the time of sending the resumes (for the full description of the validity check, see Appendix C).

Scale of discrimination

My first research question is whether there are differences in employers' reactions to white and black applicants. However, there could also be motives other than racism why one applicant has received a positive response and the other has not (Bovenkerk et al., 1995; McGinnity et al., 2009; Riach & Rich, 2002). It could be the case that an employer preferred the candidate's place of residence or internship. Any random variable could potentially result in unequal responses. However, it can be expected that these random unequal responses would arise at a similar level for both white and black applicants. Therefore, the random unequal treatment can be eliminated by deducting the unequal responses against the white candidate from the unequal responses against the black candidate. This provides the study with a measure of unequal treatment arising out of systematic actions, also known as the "net

[18] I keep a significance level of five percent throughout the study, as this is generally considered the default cut-off point (Bryman, 2016; Field, 2018).

discrimination rate".

Even though the net discrimination rate is frequently used in field experiments to measure the magnitude of discrimination, this measure is not without controversy (McGinnity et al., 2009). The problem is that it is not obvious which denominator is the most suitable for this concept. Should net discrimination be measured as a percentage of all cases that received a reaction (so also cases in which both candidates were rejected) or the cases where at least one applicant received a positive response? In this study, I have followed the tradition of other researchers and applied the second approach (Bovenkerk et al., 1995; Fibbi et al., 2021; McGinnity et al., 2009; Riach & Rich, 2002). However, when this study reported a significant finding, I also included the relative callback rate, which indicates how much more likely the white candidate is to receive a positive response compared to the black one (Fibbi et al., 2021). The advantage of this measure is that it is independent of the denominator (McGinnity et al., 2009).

Test of the hypotheses

There were 94 dyads (N) in this study when only considering those pairs in which at least one candidate received a positive response (see Table 1). Equal treatment, where both candidates were invited, occurred in most cases, namely in 74.5% (100/94*70) of the usable observations. Nonetheless, the importance of skin colour becomes vividly obvious when looking at the unequal treatment of black and white applicants side by side (See Figure 1). Unequal responses towards black candidates happened in 23.4% (100/94*22) of the usable dyads (only the white candidate received a positive response, while the black candidate did not). Conversely, unequal responses toward the white candidate appeared in only 2.1% (100/94*2) of the cases.

The unequal reactions against the black candidate minus that of the white candidate provide the net discrimination rate of 21.3% (100/94*20). In concrete terms, this rate suggests

that when equally qualified men apply for a job in the Dutch labour market, there appears to be a 21.3% chance that a black candidate will not get a callback for a job interview while a white candidate does. The net discrimination rate of 21.3% is significant at the five percent level, thereby accepting Hypothesis 1 (for the detailed computation of this hypothesis, see Appendix D). Thus, this study found that black applicants receive significantly fewer callbacks than white applicants during job applications within the Dutch labour market.

Because the difference in callbacks between white and black candidates is significant, I also calculated the relative callback rate. This study found a relative callback rate of 1.28 against black male candidates[19]. This means that to get a positive response for a job interview, black applicants must send 28 percent more applications than white applicants.

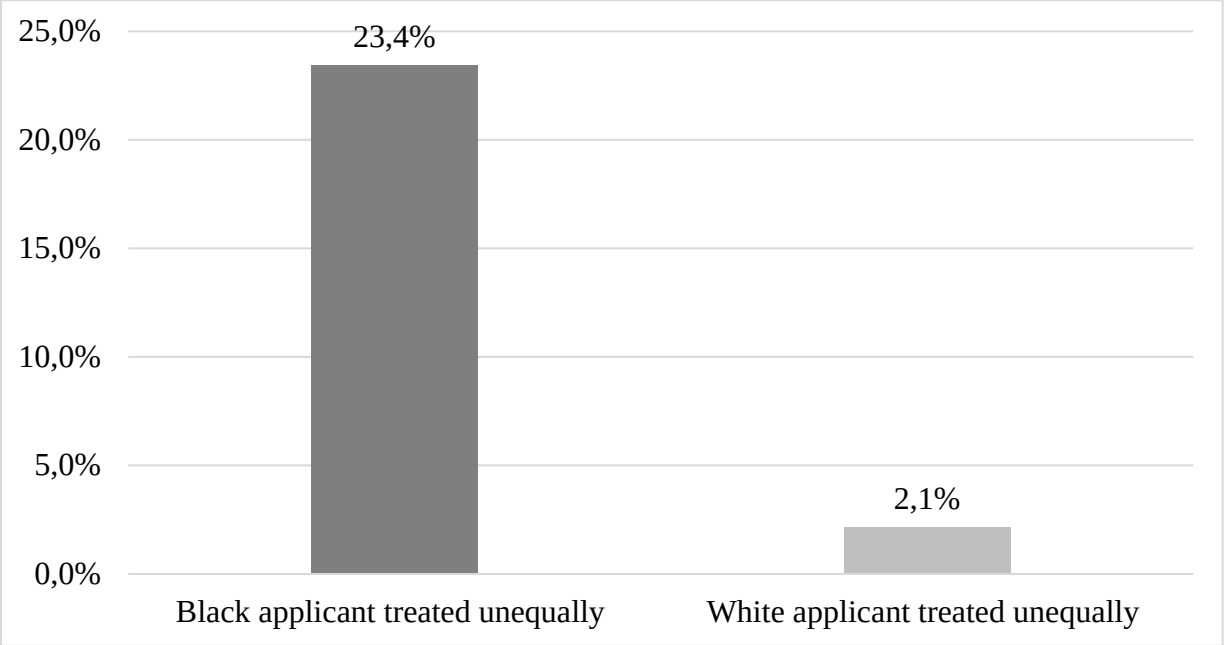
Table 1: Results of Correspondence Testing

Total pairs of applications	150
No response	14
Neither invited	42
Usable applications	94
Both invited	70
Only white applicant invited	22
Only black applicant invited	2
Net discrimination against black applicant	20
Net discrimination against black applicant in % (100/94*20)	21,3%*

Note. *p < .05

[19] Overall, the white candidate had a positive response rate of 67.7% (100/136*92), while the black candidate had a positive response rate of 52.9% (100/136*72). This translates into a relative callback rate of 1.28 (67.7%/52.9%).

Figure 1: The Difference in Unequal Treatment



In Hypothesis 2, I proposed that the difference in callbacks between white and black candidates would be smaller when applying for jobs requiring a college education than for semi-skilled jobs within the Dutch labour market. To test this hypothesis, I used a Chi-square test. The Chi-square in this study was below the threshold to reject the null hypothesis at the five percent level (see Appendix E). Therefore, I did not find a significant difference in the net discrimination rate between semi-skilled jobs and jobs requiring a college education. So, Hypothesis 2 is rejected. This conclusion was not unanticipated because there is little difference in the net discrimination rate between the education requirements of the jobs; the net discrimination rate of semi-skilled jobs is 19.6%, and for jobs requiring higher education, 22.9% (see Table 2).

Table 2: Results for Semi-Skilled Jobs and Jobs Requiring a College Education

	Semi-skilled jobs	Jobs requiring a college education
Total pairs of applications	75	75
No response	6	8
Neither invited	23	19
Usable applications	46	48
Both invited	33	37
Only white applicant invited	11	11
Only black applicant invited	2	0
Net discrimination against black applicant	9	11
Net discrimination against black applicant in %	19,6%	22,9%

Additional analysis

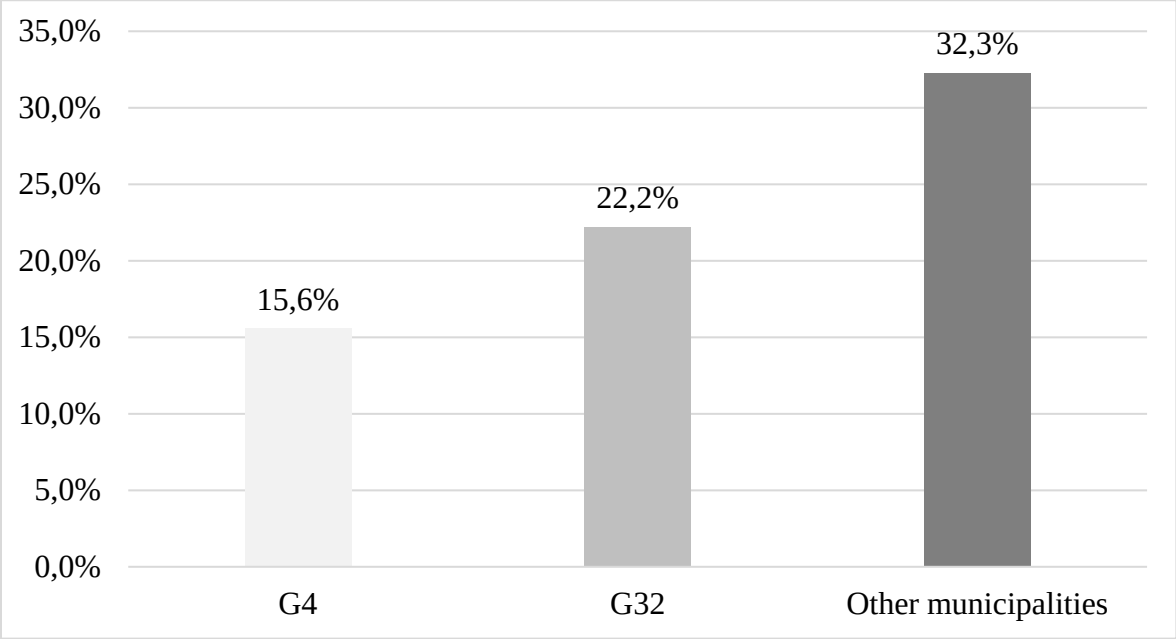
To fully utilise the data, I have conducted additional analyses. Specifically, I examined whether there were differences in the discrimination between job types and between different job locations. As in Hypothesis 2, I used a Chi-square test to assess these differences[20].

Organisations were tested in various locations. As explained in the methods section, I have distinguished between more and less urbanised areas. In this case, "G4" is the most urbanised and "other municipalities" the least. As is shown in Figure 2, there is a slight difference in the discrimination rate between the three locations in the direction that the less urbanised the municipalities, the more discrimination. However, according to the Pearson Chi-square test, these differences were not statistically significant at the five percent level (p = 0.47) (for the detailed breakdown, see Appendix F)[21].

[20] However, this time I performed the calculations with SPSS because there was a chance that an assumption of the Chi-square test would be violated. The conventional rule of thumb is that all expected values in each cell are greater than five (Field, 2018). When more than 20% of the cells have an expected value of less than five, an alternative test should be performed, such as a Fisher's exact test.

[21] In this case, I could use the Pearson Chi-square test because the number of cells below five was 16.7% and did not exceed the 20% cut-off.

Figure 2: The Net Discrimination Rate Based on Job Location



Six different job types have been tested among two distinct educational levels. I tested the jobs as financial administrator, receptionist, and sales assistant for semi-skilled jobs. In contrast, for jobs requiring a college education, I examined jobs in the positions of account management, finance, and human resources. Discrimination against the black candidate occurred in each occupation at various levels. Nevertheless, the differences between jobs were not statistically significant at the five percent level according to the Fisher exact test (p = 0.66) (for the detailed breakdown, see Appendix G)[22].

Discussion

This study focuses on skin colour as a proxy for race and examines its influence on job opportunities in the Dutch labour market. Applying for entry-level jobs with equivalent resumes and cover letters provided clear evidence of racism against black male applicants in the Netherlands. I demonstrated that black applicants received significantly fewer callbacks than white applicants during job applications. Specifically, I showed that blacks must submit

[22] 33.3% of the cells had an expected count less than five; this violated the assumption of the 20% threshold for a Chi-square test. Therefore, I had to resort to the Fisher exact test.

28 percent more job applications than equally qualified white applicants to receive a callback from employers. This rate can be considered a minimum percentage of racial discrimination against black candidates because it refers only to the callbacks for an interview and not to any possible subsequent unfair treatment (Bovenkerk et al., 1995). These racial disparities demonstrate the significance of skin colour in the Dutch labour market. Moreover, these findings challenge the just world hypothesis and contradict that every person has equal opportunities.

The differences between white and black applicants in this research were perhaps predictable based on empirical findings in other contexts. However, to the best of my knowledge, this is the first study to offer causal support for the negative effect of skin colour on job opportunities for black male candidates in the Netherlands. Therefore, I compare the results with the findings of studies from a different context. Remarkably, the relative callback rate of 1.28 found in this study is largely comparable to findings from meta-analyses in the United States (1.36) (Fibbi et al., 2021; Quillian et al., 2017). So, it appears that there are similar rates of racial discrimination, yet with a different awareness of racial problems. Whether the results of this study will translate into more awareness of racial issues in the Netherlands remains to be seen.

In this study, the prevalence of discrimination against blacks became obvious when comparing the results of the applications side-by-side. Without this comparison, there would be no direct evidence of discrimination because the employers did not display any explicit bias against the fictitious black candidate. For example, none of the employers stated that the candidate's skin colour had something to do with their rejection. Thus, the discrimination against black candidates in this study is not characterised by overt racism or clear prejudice expressed towards the candidates. Instead, the behaviour of employers is instrumental in demonstrating that black job applicants are discriminated against in the labour market. This

observation is also in line with the ideas of van den Broek (2014) and Essed (1991) that racism has taken more covert and subtle forms in the Netherlands. It might also explain why skin colour or racial issues are a low-key topic in Dutch discourse.

However, this study demonstrated that even in the 21st century, skin colour has real consequences on a person's life chances and that the Netherlands is not "innocent and inherently non-racist". Moreover, I showed that a more nuanced political and academic discourse is needed, where skin colour and race become part of the dialogue rather than merely referring to presumably innocent identities (ethnicity, minority, or autochtoon/allochtoon). Lastly, this study proved that Essed's (1991) call for more research into racism within the Netherlands is still relevant today.

Nevertheless, this study did not find significant evidence for Hypothesis 2, which stated that the difference in callbacks between white and black candidates would be smaller when applying for jobs requiring a college education than semi-skilled jobs. Therefore, I did not find that a higher-level educational degree decreased racial discrimination against blacks. This result contradicts several previous findings within the literature (Bertrand & Mullainathan, 2004; Bovenkerk et al., 1995; Flage, 2019). However, it is necessary to note that earlier studies have used semi-skilled jobs as jobs that do not require any education (Bovenkerk et al., 1995; Flage, 2019). In contrast, jobs requiring a college education typically involved a scientific degree. In this study, both job levels held a vocational education, which might be seen as similar in the eyes of employers. This could explain why this study did not find a significant difference between the different educational requirements of the jobs. Another explanation of this finding might be that there is no difference in racial discrimination between different levels of education in the Netherlands. Employers might focus on physical appearance attributes regardless of the job positions or tightness of the market. Moreover, educational qualifications may not be enough to counteract the generalised

negative perceptions of racial minorities. The last explanation for the insignificance could be due to methodological issues that made it unlikely to find significant results. I will elaborate on this in the limitations. Clearly, more work needs to be done to understand the interaction between education and skin colour and its effect on job opportunities in the Netherlands.

This study hinted at an interesting possible link in the additional analyses. However, there needs to be more research to draw any conclusions. In the explorations, I looked at whether there were different levels of racial discrimination within different geographical contexts. It seemed as if the less urbanised the municipalities, the more racial discrimination against the fictitious black candidate occurred. Even though the results pointed in this direction, there was not enough power to be considered statistically significant due to the relatively small sample size. However, the small sample might have also caused a skewed presentation of the findings, so this suggestion should be treated with caution. The potential differences in discrimination rates between urban and rural settings form an interesting area for future research. Especially since research in other contexts, such as Switzerland, has found significant differences in the discrimination rates between rural and urban settings (Fibbi et al., 2021).

Thus, I showed that employers discriminate against black male applicants in the Dutch labour market. In addition, I did not find that a higher-level educational degree would reduce discrimination. Even though this study is not focused on policy implementations, my insights suggest that policymakers should pay more attention to employers rather than interventions aimed at improving the education of racial minorities to enhance employment opportunities in the Netherlands. Although there is no consensus in the literature on how to deal with employment discrimination, based on my findings, it seems important to emphasise standardising hiring procedures, pre-determining qualifications, anonymous applications, and clear and concrete anti-discrimination policies by the Dutch government (Bragger et al., 2002;

Fang et al., 2018; French, 2017; Krause et al., 2012; Midtbøen, 2014).

Despite the implications discussed above, the current research has several limitations. I will start with one of the key limitations, namely that this study focused specifically on male candidates. Because this study did not assess how employers would respond to women (or other gender identities), the findings cannot be generalised to all black candidates. The discrimination rate might have varied if the tests were conducted for female applicants. Since discrimination in the labour market can have destructive effects on people's lives, it is essential that research extensively explores the potential effects for every person that could be impacted. Having clear insight into all systemic mechanisms at play and their interaction is vital to breaking these systems down. Policymakers need to be thoroughly informed if they aim to protect marginalised groups against structural discrimination in the future. Therefore, I want to underline that future research should study the extent skin colour influences hiring chances for women and other gender identities.

The following limitations are concerned with the methods of this study. By their nature, correspondence tests only examine a particular point in the application process. Specifically, this method only measures the callbacks for interviews and does not care about whether the applicants get the position and the proposed salary conditional on getting the job. Reduced invitations to job interviews would likely translate into reduced job offers. Nevertheless, based on these results, I cannot make any claims about gaps in hiring and earnings. This is not necessarily a serious issue because evidence suggests that discrimination mainly occurs during the selection for interviews when information about the candidate is at a minimum, not at the later stages (Bovenkerk, 1992; Pager, 2007). Moreover, social and personal networks are a common path by which people find jobs, but these channels could not be studied with the current methodology. Further, the photos used in this study were kept as similar as possible. However, I have not empirically tested the perceived attractiveness,

competence, age, and warmth of the people in these pictures. Hence, any of these factors could have affected the results.

Another critical limitation is the size of the sample. All data was collected as part of a thesis, for which there were few resources. This meant I had to send out all applications manually within a limited time frame. Compared to studies from several decades ago, the sample size of this study is sufficient (e.g., Jolson, 1974; McGinnity et al., 2009; Pager, 2003), but compared to recent studies, it is relatively small (e.g., Campos-Vazquez & Gonzalez, 2020; Thijsen et al., 2020). The relatively small sample could have decreased statistical power to detect effects (Acock, 2005; Field, 2018). Moreover, a small sample can be problematic because findings tend to be incorrect estimates of the corresponding population (Leppink et al., 2016). This means that this research may form a conservative test of the hypotheses. Instead, using a larger data set would permit more powerful hypothesis tests and control for other variables. As such, future research should investigate a larger sample with more dyads. This would also allow more characteristics and intersections to be examined alongside skin colour and the effect on job opportunities.

The following limitation is the scope of this study. The effects of racism might vary across the population, perhaps concentrated in certain groups (Pager et al., 2009). Yet, many important aspects were not examined in this study. With all experiments of this nature, it was not feasible to explore all segments of the labour market, such as the public sector (e.g., education, healthcare) and jobs requiring years of experience (e.g., directors, doctors) (McGinnty et al., 2009). Moreover, this study was only focused on a few job positions. Even though I did not find a significant difference in the discrimination rate between jobs, future research should explore this more thoroughly. In addition, there are numerous intersections that, in combination with skin colour, could affect labour market opportunities (e.g., family situation, handicap, sexuality, and religion). Therefore, future research is encouraged to

examine whether and to what extent these characteristics combined with skin colour would influence hiring chances in the Dutch labour market.

Lastly, racism is often concealed and is not a straightforward matter to examine (Pager, 2003). The experiment in this study allows me to infer the effect of skin colour on labour market opportunities with great certainty, but the effects of racism are narrowly defined. For example, the broader impacts of racism over the life course are more difficult to establish (Pager et al., 2009). In addition, it is not evident why employers have discriminated against the fictitious applicants. For this reason, future research should employ triangulation, that is, a combination of methods to try to expose a more comprehensive understanding of racial discrimination.

In conclusion, this study is important in the broader context of the just world hypothesis and the continuation of systemic racism. I realise that the observations in this study are just one aspect of racism, so the findings that I have reported are only a small part of a larger issue. Yet, I was able to provide a clear and measurable indicator of the potential racial discrimination against black men and their likelihood of getting a job interview in the Dutch labour market. This paper will hopefully stimulate future research into the role of skin colour and race in the Dutch context to create a society where everyone has equal opportunities.

References

Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family, 67*(4), 1012–1028. https://doi.org/10.1111/j.1741-3737.2005.00191.x

Adelman, L. (2014, April 24). Race: The power of an illusion [Video]. YouTube. https://www.youtube.com/watch?v=Y8MS6zubIaQ

Amiraux, V., & Simon, P. (2006). There are no minorities here. International Journal of Comparative Sociology, 47(3–4), 191–215. https://doi.org/10.1177/0020715206066164

Andriessen, I., Nievers, E., Dagevos, J., & Faulk, L. (2012). Ethnic discrimination in the Dutch labor market. Work and Occupations, 39(3), 237–269. https://doi.org/10.1177/0730888412444783

Andriessen, I., Nievers, E., Faulk, L., & Dagevos, J. (2010). Liever Mark dan Mohammed? Onderzoek naar arbeidsmarktdiscriminatie van niet-Westerse migranten via praktijktests. Sociaal En Cultureel Planbureau.

Austin, A. (2013). The unfinished march: An overview. Economic Policy Institute. https://www.epi.org/publication/unfinished-march-overview/

Baert, S. (2018). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. Audit Studies: Behind the Scenes with Theory, Method, and Nuance, 63–77. https://doi.org/10.1007/978-3-319-71153-9_3

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is there less discrimination in occupations where recruitment is difficult? ILR Review, 68(3), 467–500. https://doi.org/10.1177/0019793915570873

Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. In E. Duflo & A. Banerjee (Eds.), Handbook of economic field experiments, (pp. 309–393). North Holland.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American Economic Review, 94(4), 991–1013. https://doi.org/10.1257/0002828042002561

Biernat, M., & Kobrynowicz, D. (1997). Gender- and race-based standards of competence: Lower minimum standards but higher ability standards for devalued groups. Journal of Personality and Social Psychology, 72(3), 544–557. https://doi.org/10.1037/0022-3514.72.3.544

Blommaert, L., Coenders, M., & van Tubergen, F. (2013). Discrimination of Arabic-named applicants in the Netherlands: An internet-based field experiment examining different phases in online recruitment procedures. Social Forces, 92(3), 957–982. https://doi.org/10.1093/sf/sot124

Bobo, L. D., Charles, C. Z., Krysan, M., & Simmons, A. D. (2012). The real record on racial attitudes. In P. V. Marsden (Ed.), Social Trends in American Life: Findings from the General Social Survey Since 1972 (pp. 38–83). Princeton.

Bonilla-Silva, E. (2016). Racism without racists, color-blind racism and the persistence of racial inequality in the United States. Rowman & Littlefield.

Bovenkerk, F. (1977). Rasdiscriminatie op de Amsterdamse arbeidsmarkt. Sociologische Gids, 24(1–2), 58–75.

Bovenkerk, F., Gras, M., & Ramsoedh, D. (1995). Discrimination against migrant workers and ethnic minorities in access to employment in the Netherlands. International Labour Office.

Bovenkerk, F. (1992). *Testing discrimination in natural experiments*. International Labour Office.

Bragger, J. D., Kutcher, E., Morgan, J., & Firth, P. (2002). The effects of the structured interview on reducing biases against pregnant job applicants. Sex Roles, 46(7/8), 215–226. https://doi.org/10.1023/a:1019967231059

Bryman, A. (2016). Social research methods. Oxford University Press.

Campos-Vazquez, R. M., & Gonzalez, E. (2020). Obesity and hiring discrimination. Economics & Human Biology, 37. https://doi.org/10.1016/j.ehb.2020.100850

Cancio, A. S., Evans, T. D., & Maume, D. J. (1996). Reconsidering the declining significance of race: Racial differences in early career wages. American Sociological Review, 61(4), 541–556. https://doi.org/10.2307/2096391

Centraal Bureau voor de Statistiek. (2018). Grote gemeenten. https://www.cbs.nl/nl-nl/achtergrond/2018/09/niet-alle-naoorlogse-stadswijken-kennen-achterstand/grote-gemeenten#:%7E:text=De%20G4%20bestaat%20uit%20de,gemeenten%20tot%20de%20zogeheten%20G32

Centraal Bureau voor de Statistiek. (2020a). Gemiddeld persoonlijk inkomen 2018 en 2019. Retrieved 2021, from https://www.cbs.nl/nl-nl/maatwerk/2021/08/gemiddeld-persoonlijk-inkomen-2018-en-2019

Centraal Bureau voor de Statistiek. (2020b). Kerncijfers arbeidsdeelname. https://mvstat.cbs.nl/#/MVstat/nl/dataset/26011NED/table?dl=6604E

Centraal Bureau voor de Statistiek. (2020c). Mannen en vrouwen. https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/mannen-en-vrouwen

Centraal Bureau voor de Statistiek. (2021). Werkloosheid naar onderwijsniveau. Retrieved 2021, from https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werklozen/werkloosheid-naar-onderwijsniveau

Chrobot-Mason, D. (2004). Managing racial differences. Group & Organization Management, 29(1), 5–31. https://doi.org/10.1177/1059601103252102

de Beer, P., & van Pinxteren, M. (2016). Meritocratie: Op weg naar een nieuwe klassensamenleving? Amsterdam University Press. https://doi.org/10.5117/9789462983397

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? Personality and Social Psychology Bulletin, 21(11), 1139–1150. https://doi.org/10.1177/01461672952111002

Di Stasio, V., & Lancee, B. (2020). Understanding why employers discriminate, where and against whom: The potential of cross-national, factorial and multi-group field experiments. Research in Social Stratification and Mobility, 65, 100463. https://doi.org/10.1016/j.rssm.2019.100463

Dovidio, J. F., Gaertner, S. L., Fiske, S. T., Gilbert, D. T., & Lindzey, G. (2010). Intergroup bias. In Handbook of social psychology (pp. 1084–1121). John Wiley & Sons. https://doi.org/10.1002/9780470561119.socpsy002029

Duru-Bellat, M., & Tenret, E. (2012). Who's for meritocracy? Individual and contextual variations in the faith. Comparative Education Review, 56(2), 223–247. https://doi.org/10.1086/661290

Essed, P. (1991). Knowledge and resistance: Black women talk about racism in the Netherlands and the USA. Feminism & Psychology, 1(2), 201–219. https://doi.org/10.1177/0959353591012003

Essed, P., & Hoving, I. (2014). Innocence, smug ignorance, resentment: An introduction to Dutch racism. In P. Essed & I. Hoving (Eds.), Dutch Racism (pp. 9–30). Brill.

Fang, A. H., Guess, A. M., & Humphreys, M. (2019). Can the government deter discrimination? Evidence from a randomised intervention in new york city. The Journal of Politics, 81(1), 127–141. https://doi.org/10.1086/700107

Fibbi, R., Ruedin, D., Stünzi, R., & Zschirnt, E. (2021). Hiring discrimination on the basis of skin colour? A correspondence test in Switzerland. Journal of Ethnic and Migration Studies, 48(7), 1515–1535. https://doi.org/10.1080/1369183x.2021.1999795

Field, A. (2018). Discovering Statistics Using IBM SPSS Statistics. SAGE Publications.

Flage, A. (2019). Discrimination against gays and lesbians in hiring decisions: A meta-analysis. International Journal of Manpower, 41(6), 671–691. https://doi.org/10.1108/ijm-08-2018-0239

Frankenberg, R. (1993). White Women, Race Matters. Amsterdam University Press.

Gharib, S. E. (2022). What is intersectionality and why is it important? Global Citizen. https://www.globalcitizen.org/en/content/what-is-intersectionality-explained/#:%7E:text=This%20is%20why%20intersectionality%20is,differs%20within%20different%20social%20contexts.

Goffman, E. (1963). Stigma: Notes on the management of spoiled identity. Gardners Books.

Griffith, D. M. (2012). An intersectional approach to men's health. Journal of Men's Health, 9(2), 106–112. https://doi.org/10.1016/j.jomh.2012.03.003

Harris, A., & Leonardo, Z. (2018). Intersectionality, Race-Gender subordination, and education. Review of Research in Education, 42(1), 1–27. https://doi.org/10.3102/0091732x18759071

Homan, A. C., van Knippenberg, D., van Kleef, G. A., & de Dreu, C. K. W. (2007). Bridging faultlines by valuing diversity: Diversity beliefs, information elaboration, and performance in diverse work groups. Journal of Applied Psychology, 92(5), 1189–1199. https://doi.org/10.1037/0021-9010.92.5.1189

Hondius, E. G. (2014). Black Dutch voices: Reports from a country that leaves racism unchallenged. In P. Essed & I. Hoving (Eds.), Dutch Racism (pp. 273–293). Brill.

Jolson, M. A. (1974). Employment barriers in marketing. Journal of Marketing, 38(2), 67. https://doi.org/10.2307/1250200

Jost, J., & Hunyady, O. (2003). The psychology of system justification and the palliative function of ideology. European Review of Social Psychology, 13(1), 111–153. https://doi.org/10.1080/10463280240000046

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened résumés. Administrative Science Quarterly, 61(3), 469–502. https://doi.org/10.1177/0001839216639577

Kerckhoff, A. C., Raudenbush, S. W., & Glennie, E. (2001). Education, cognitive skill, and labor force outcomes. Sociology of Education, 74(1), 1. https://doi.org/10.2307/2673142

Kinder, D. R., & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. Journal of Personality and Social Psychology, 40(3), 414–431. https://doi.org/10.1037/0022-3514.40.3.414

Krause, A., Rinne, U., & Zimmermann, K. F. (2012). Anonymous job applications of fresh Ph.D. economists. Economics Letters, 117(2), 441–444. https://doi.org/10.1016/j.econlet.2012.06.029

Lancee, B. (2019). Ethnic discrimination in hiring: Comparing groups across contexts. Results from a cross-national field experiment. Journal of Ethnic and Migration Studies, 47(6), 1181–1200. https://doi.org/10.1080/1369183x.2019.1622744

Laveist, T. A. (1996). Why we should continue to study race. . . but do a better job: An essay on race, racism and health. Ethnicity and Disease, 21–29.

Leppink, J., Winston, K., & O'Sullivan, P. (2016). Statistical significance does not imply a real effect. *Perspectives on Medical Education*, *5*(2), 122–124. https://doi.org/10.1007/s40037-016-0256-6

Lerner, M. (1980). The belief in a just world: A fundamental delusion. Plenum.

List, J. (2009). Informed consent in social science response. Science 322(5902), 672.

Luijkx, R., & Wolbers, M. H. J. (2009). The effects of non-employment in early work-life on subsequent employment chances of individuals in the Netherlands. European Sociological Review, 25(6), 647–660. https://doi.org/10.1093/esr/jcp002

McGinnity, F., Nelson, J., Lunn, P., & Quinn, E. (2009). Discrimination in recruitment: Evidence from a field experiment. The Equality Authority and The Economic and Social Research Institute.

Midtbøen, A. H. (2014). The context of employment discrimination: Interpreting the findings of a field experiment. The British Journal of Sociology, 66(1), 193–214. https://doi.org/10.1111/1468-4446.12098

Ministerie van Onderwijs, Cultuur en Wetenschap. (2021, August). Hoogst behaald onderwijsniveau. Onderwijs Algemeen - OCW in cijfers. https://www.ocwincijfers.nl/sectoren/onderwijs-algemeen/hoogst-behaald-onderwijsniveau

Moss, P., & Tilly, C. (2003). Stories Employers Tell. SAGE Publications.

Nimako, K., & Willemsen, G. (1993). Multiculturalisme, verzuilde, samenleving and verzorgingsstaat: Naar een pluralistische democratie. Achter de Coulissen: Gedachten over de Multi-Etnische Samenleving.

O'Donnell, S., & Richardson, N. (2020). No country for Middle-Aged men? International Journal of Mens Social and Community Health, 3(2), e32–e45. https://doi.org/10.22374/ijmsch.v3i2.32

Okazaki, S., & Sue, S. (2016). Methodological issues in assessment research with ethnic minorities. Methodological Issues and Strategies in Clinical Research (4th Ed.)., 235–247. https://doi.org/10.1037/14805-015

Omi, B. M., & Winant, H. (2014). Racial formation in the United States. Routledge. https://doi.org/10.4324/9780203076804

Pager, D. (2003). The mark of a criminal record. American Journal of Sociology, 108(5), 937–975. https://doi.org/10.1086/374403

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. The American Academy of Political and Social Science, 609(1), 104–133. https://doi.org/10.1177/0002716206294796

Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a Low-Wage labor market. American Sociological Review, 74(5), 777–799. https://doi.org/10.1177/000312240907400505

Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. Annual Review of Sociology, 34(1), 181–209. https://doi.org/10.1146/annurev.soc.33.040406.131740

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. Proceedings of the National Academy of Sciences, 114(41), 10870–10875. https://doi.org/10.1073/pnas.1706255114

Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. The Economic Journal, 112(483), 480–518. https://doi.org/10.1111/1468-0297.00080

Sabharwal, M. (2014). Is diversity management sufficient? Organisational inclusion to further performance. Public Personnel Management, 43(2), 197–217. https://doi.org/10.1177/0091026014522202

Sesardic, N. (2010). Race: A social destruction of a biological concept. Biology & Philosophy, 25(2), 143–162. https://doi.org/10.1007/s10539-009-9193-7

Smith, A. (2020). The internet and job seeking. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2015/11/19/1-the-internet-and-job-seeking/

Sociaal en Cultureel Planbureau. (2010). Discriminatiemonitor nietwesterse migranten op de arbeidsmarkt 2010. Ministerie van Sociale Zaken En Werkgelegenheid.

Steijn, S., van de Werfhorst, H., & Burgoon, B. (2016). Oordelen over het meritocratische gehalte van de Nederlandse samenleving. In P. de Beer & M. van Pinxteren (Eds.), Meritocratie: op weg naar een nieuwe klassensamenleving? (pp. 199–207). Amsterdam University Press. https://doi.org/10.5117/9789462983397

Thijssen, L., Coenders, M., & Lancee, B. (2020). Ethnic discrimination in the Dutch labor market: Differences between ethnic minority groups and the role of personal information about job Applicants—Evidence from a field experiment. Journal of International Migration and Integration, 22(3), 1125–1150. https://doi.org/10.1007/s12134-020-00795-w

van den Broek, L. M. (2012). De ironie van gelijkheid. Kantharos.

van den Broek, L. M. (2014). Neither with, nor without them, ethnic diversity on the work floor: How egalitarianism breeds discrimination. Dutch Racism, 255–271. https://doi.org/10.1163/9789401210096_014

van Knippenberg, D., de Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. Journal of Applied Psychology, 89(6), 1008–1022. https://doi.org/10.1037/0021-9010.89.6.1008

van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. Annual Review of Psychology, 58(1), 515–541. https://doi.org/10.1146/annurev.psych.58.110405.085546

Veenman, J. (2010). Measuring labor market discrimination: An overview of methods and their characteristics. American Behavioral Scientist, 53(12), 1806–1823. https://doi.org/10.1177/0002764210368098

Warner, D. F., & Brown, T. H. (2011). Understanding how race/ethnicity and gender define age-trajectories of disability: An intersectionality approach. Social Science & Medicine, 72(8), 1236–1248. https://doi.org/10.1016/j.socscimed.2011.02.034

Weinberg, D., Stevens, G. W., Currie, C., Delaruelle, K., Dierckens, M., Lenzi, M., Main, G., & Finkenauer, C. (2021). Country-Level meritocratic beliefs moderate the social gradient in adolescent mental health: A multilevel study in 30 European countries. Journal of Adolescent Health, 68(3), 548–557. https://doi.org/10.1016/j.jadohealth.2020.06.031

Wekker, G. (2016). White innocence. Amsterdam University Press.

Wekker, G., & Lutz, H. (2001). Een hoogvlakte met koude winden: De geschiedenis van het gender-en etniciteitsdenken in Nederland. Caleidoscopische Visies: Zwarte, Migranten-En Vluchtelingen Vrouwenbeweging in Nederland.

Western Washington University. (2014). Chi-square notes. https://fire.biol.wwu.edu//trent/trent/chisquarenotes.pdf

Williams, J., & Wilson, V. (2019). Black workers endure persistent racial disparities in employment outcomes. Economic Policy Institute. https://www.epi.org/publication/labor-day-2019-racial-disparities-in-employment/

Appendix

Appendix A: Explanation of the Selected Jobs

Bel Hadi's first advice was to apply for jobs in the private sector because of the current labour market situation. Bel Hadi pointed out significant shortages in public sectors such as education and health care. Almost everyone would be hired in these sectors, even if applicants would not have the right work experience. Instead, he advised conducting research in business-related fields.

To determine whether there was a difference in candidates' chances of success, the candidates needed a reasonable chance of being invited for a job interview. Yet, producing CVs and letters for each vacancy was not possible within resource and time constraints. Therefore, Bel Hadi advised focusing on a limited number of jobs. However, he also indicated that focusing on one or two specific jobs would be problematic due to the limited time frame of this study. Instead, Bel Hadi argued that it would be more sensible to use a small selection of business-related jobs.

Appendix B: Explanation of the Procedures

As shown in the table below, the order in which the resumes (and related attributes) were submitted was rotated during the data collection. For instance, first I applied by submitting “CV one” with a photo of a black person and “cover letter one”. A couple of hours later, I applied for the same job by sending “CV two” with a picture of a white person and “cover letter two”. The order was then changed when I reacted to the next job vacancy. This procedure was followed throughout the data collection.

Table: Order of Applications

Job Vacancy	First Application			Second Application		
	CV	Photo of Applicant	Cover Letter	CV	Photo of Applicant	Cover Letter
1	CV 1	Black	Letter 1	CV 2	White	Letter 2
2	CV 1	White	Letter 1	CV 2	Black	Letter 2
3	CV 2	White	Letter 1	CV 1	Black	Letter 2
4	CV 2	Black	Letter 1	CV 1	White	Letter 2
5	CV 2	White	Letter 2	CV 1	Black	Letter 1
6	CV 2	Black	Letter 2	CV 1	White	Letter 1
7	CV 1	Black	Letter 2	CV 2	White	Letter 1
8	CV 1	White	Letter 2	CV 2	Black	Letter 1

I responded to publicly advertised vacancies on online job markets (e.g., job portals) because this is the most common way of applying for jobs nowadays (Smith, 2020). Positions advertised by recruiters, instead of employers themselves, were excluded from the study to minimise the risk of detection (McGinnity et al., 2009). In addition, some job positions presented technical obstacles that prevented them from being examined. For example, I eliminated vacancies listed on websites with IP scanners, checked home addresses, and required extensive cover letters. Lastly, an employer or organisation could only be included once in the data; those where I had previously applied were excluded from subsequent applications.

Appendix C: Validity Check

It is essential to determine whether one of the resumes or cover letters got significantly more positive responses (Baert, 2018; Bertrand & Duflo, 2017; Bovenkerk et al., 1995). This could happen when one of the resumes or letters was unconsciously more appealing or had higher quality. The following validity check was applied to examine these potential differences. Overall, there were 24 unequal responses where one applicant received a positive reaction and the other did. Unequal treatment occurred 12 times when "resume one" was used and 12 times when "resume two" was applied[23]. Moreover, in 11 cases "cover letter one" was utilised and in 13 cases "letter two" was employed.

The null hypothesis is that the cover letter or resume is not the cause of the unequal response. The likelihood of a positive reaction is than 50% for both resumes and cover letters. In an ideal situation, this would mean that of the 24 unequal responses, 12 cases should be due to "cover letter one" and "resume one", and 12 cases should emerge from "cover letter two" and "resume two". Following Bovenkerk et al. (1995), I used the standard deviation of the binomial curve F = $\sqrt{npq}$ to compute the deviation for which the null hypothesis is not rejected at the five percent level. In this case, p and q are the probability of positive reactions to "cover letter or resume one" and "cover letter or resume two". So, F = $\sqrt{24 * 0.5 * 0.5}$ = 2.45. The null hypothesis at the five percent level will not be rejected when the mean value is not more than ± 1.96F in a normal distribution. Put differently, a resume or cover letter may not contain more or less than 12 ± 4 = less than eight or more than 16 cases out of the 24 total unequal responses. The resumes exhibited a perfect distribution. Both resumes had the exact same likelihood of an unequal response, which leads to the conclusion that unequal treatment

[23] Even though the resumes differed in format between candidates applying for semi-skilled jobs and jobs requiring a college education, I did bundle them together as one group for the validity check. The rationale behind this decision is that the content of the resumes did not differ, apart from the level of education. Bundling the two types of resumes together also provides the validity check with more statistical power.

is not due to the resumes. Moreover, because the unequal reactions to the various cover letters were 11 and 13 respectively, it can be concluded that unequal treatment was not caused by the letters.

Furthermore, it is relevant to exclude that the time of sending the applications did not affect unequal treatment. Of the 24 cases where unequal responses appeared, 12 occurred when the application was sent first and 12 when the application was sent second. Because there should be a 50% chance of unequal treatment among these different timings, and this distribution was precisely the case in this study, it can also be ruled out that the timing of the applications caused unequal treatment.

Appendix D: Testing Hypothesis 1

There were 94 dyads (N) in this study when only considering those pairs in which at least one candidate received a positive response. For this sample, I must compute the minimum net discrimination rate needed to reject the null hypothesis at the five percent threshold. Put differently, when can I accept Hypothesis 1 and assume that black applicants receive fewer callbacks than white applicants during job applications within the Dutch labour market.

Once more, I followed the method of Bovenkerk et al. (1995) by applying the formula: $\sqrt{N} = 1.96F/CR*F$. In this case, N is the number of usable dyads, 1.96 is the z-value at the five percent significance level of the normal distribution, F is the standard value of the normal distribution, and CR is the critical net discrimination rate. Thus, for N = 94, the minimum net discrimination rate to assume that employers have racially discriminated against the applicants at the five percent significance level is $CR = 1.96F/\sqrt{N}*F = 1.96/\sqrt{94} = 20.2\%$. In other words, the null hypothesis can be rejected when the net discrimination rate is higher than the critical rate of 20.2%.

This study had a net discrimination rate of 21.3% (100/94*20) against black applicants. The discrimination rate of 21.3% is higher than the 20.2% needed to reject the null hypothesis at the five percent level, so black applicants receive significantly fewer callbacks than white applicants.

Appendix E: Testing hypothesis 2

To conclude that there is a significant difference at the five percent level, the value of Chi-square for the difference in net discrimination needs to be larger than 3.84 with the degrees of freedom of one (*df*=1)[24]. To calculate the Chi-square, I used the formula:

$$x^2 = \sum \frac{(observed - expected)^2}{expected}$$

The expected value is based on the null hypothesis that there is no difference in net discrimination between semi-skilled jobs and jobs requiring a college education. This implies that the net discrimination should be evenly distributed between the two categories. Specifically, the expected value = usable applications of a category * the general net discrimination rate against black applicants in percentages (without considering the distribution of the categories as presented in Table 1). In other words, the expected value of net discrimination for semi-skilled jobs is (46*21.3% =) 9.798 and for jobs requiring a college education, (48*21.3% =) 10.224 (see Table 2)[25] [26]. This results in the following equation:

$$x^2 = \frac{(9-9.798)^2}{9.798} + \frac{(11-10.224)^2}{10.224} = 0.124.$$

The Chi-square 0.124 is below the threshold of 3.84 to reject the null hypothesis. Hence, there was no significant difference in the net discrimination rate between semi-skilled jobs and jobs requiring a college education at the five percent level.

[24] I used *df* = 1, because there are two categories (semi-skilled jobs and jobs requiring a college education) and *df* = number of categories – 1.
[25] To avoid rounding errors, expected values in Chi-square calculations should be expressed to four decimal places (Western Washington University, 2014).
[26] The Chi-square test demands that at least 20% of the categories have an expected count greater than five, otherwise the test will be invalid (Field, 2018). Both categories have an expected value above five, so this assumption is not violated.

Appendix F: Description of the Chi-Square Test for Different Job Locations

Table: Results of Correspondence Testing Based on Job Location

	G4	G32	Other municipalities
Total pairs of applications	66	30	54
No response	7	4	3
Neither invited	14	8	20
Usable applications	45	18	31
Both invited	36	13	21
Only white applicant invited	1	1	0
Only black applicant invited	8	4	10
Net discrimination against black applicant	7	4	10
Net discrimination against black applicant in %	15,6%	22,2%	32,3%

Table: SPSS Computation

		Equal treatment	Unequal treatment	Total
G4	Count	36	9	45
	Expected count	33,5	11,5	45
G32	Count	13	5	18
	Expected count	13,4	4,6	18
Other municipalities	Count	21	10	31
	Expected count	23,1	7,9	31
Total	Count	70	24	94
	Expected count	70	24	94

Note. One cell (16.7%) has an expected count of less than five. Pearson Chi-square $p = 0.47$. Unequal treatment happened when one applicant received a positive response and the other did not.

Appendix G: Description of the Chi-Square Test for Different Job Types

Table: Results of Correspondence Testing for Different Occupations

Semi-skilled jobs			
	Financial administrator	Receptionist	Sales assistant
Total pairs of applications	35	20	20
No response	2	3	1
Neither invited	11	8	4
Usable applications	22	9	15
Both invited	14	7	12
Only white applicant invited	1	0	1
Only black applicant invited	7	2	2
Net discrimination against black applicant	6	2	1
Net discrimination against black applicant in %	27,3%	22,2%	6,7%
Jobs requiring a college education			
	Account manager	HR-employee	Finance employee
Total pairs of applications	35	20	20
No response	3	2	3
Neither invited	11	2	6
Usable applications	21	16	11
Both invited	16	11	10
Only white applicant invited	0	0	0
Only black applicant invited	5	5	1
Net discrimination against black applicant	5	5	1
Net discrimination against black applicant in %	23,8%	31,3%	9,1%

Table: SPSS Computation

		Equal treatment	Unequal treatment	Total
Financial administrator	Count	14	8	22
	Expected count	16,4	5,6	22
Receptionist	Count	7	2	9
	Expected count	6,7	2,3	9
Sales assistant	Count	12	3	15
	Expected count	11,2	3,8	15
Account manager	Count	16	5	21
	Expected count	15,6	5,4	21
HR-employee	Count	11	5	16
	Expected count	11,9	4,1	16
Finance employee	Count	10	1	11
	Expected count	8,2	2,8	11
Total	Count	70	24	94
	Expected count	70	24	94

Note. Four cells (33.3%) have an expected count of less than five. Therefore, the Pearson Chi-square is invalid. Instead, Fisher's exact test $p = 0.66$. Unequal treatment happened when one applicant received a positive response and the other did not.

Appendix H: Ethics and Privacy Checklist

PART I: GENERAL INFORMATION

Project title: *Discrimination in the Dutch Labour Market: A Correspondence Test of Race and Education*

Name, email of student: *Dennis van Doornik - 625516dd@student.eur.nl*

Name, email of supervisor: *Dr Bonnie French - french@essb.eur.nl (second: Jennifer Holland - j.a.holland@essb.eur.nl)*

Start date and duration: *Jan- Jun 2022*

Is the research study conducted within DPAS **YES** - NO

If 'NO': at or for what institute or organization will the study be conducted?
(e.g. internship organization)

PART II: HUMAN SUBJECTS

1. Does your research involve human participants. **YES** - NO

   If 'NO': skip to part V.

   If 'YES': does the study involve medical or physical research? YES - **NO**
   Research that falls under the Medical Research Involving Human Subjects Act (WMO) must first be submitted to an accredited medical research ethics committee or the Central Committee on Research Involving Human Subjects (CCMO).

2. Does your research involve field observations without manipulations that will not involve identification of participants. YES - **NO**

   If 'YES': skip to part IV.

3. Research involving completely anonymous data files (secondary data that has been anonymized by someone else). YES - **NO**

   If 'YES': skip to part IV.

PART III: PARTICIPANTS

1.	Will information about the nature of the study and about what participants can expect during the study be withheld from them?	**YES** - NO
2.	Will any of the participants not be asked for verbal or written 'informed consent,' whereby they agree to participate in the study?	**YES** - NO
3.	Will information about the possibility to discontinue the participation at any time be withheld from participants?	**YES** - NO
4.	Will the study involve actively deceiving the participants? *Note: almost all research studies involve some kind of deception of participants. Try to think about what types of deception are ethical or non-ethical (e.g. purpose of the study is not told, coercion is exerted on participants, giving participants the feeling that they harm other people by making certain decisions, etc.).*	**YES** - NO
5.	Does the study involve the risk of causing psychological stress or negative emotions beyond those normally encountered by participants?	YES - **NO**
6.	Will information be collected about special categories of data, as defined by the GDPR (e.g. racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for the purpose of uniquely identifying a person, data concerning mental or physical health, data concerning a person's sex life or sexual orientation)?	YES - **NO**
7.	Will the study involve the participation of minors (<18 years old) or other groups that cannot give consent?	YES - **NO**
8.	Is the health and/or safety of participants at risk during the study?	YES - **NO**
9.	Can participants be identified by the study results or can the confidentiality of the participants' identity not be ensured?	YES - **NO**
10.	Are there any other possible ethical issues with regard to this study?	**YES** - NO

If you have answered 'YES' to any of the previous questions, please indicate below why this issue is unavoidable in this study.

I have discussed the ethical considerations at length in the methods. Although this study technically involves human participants, the interactions I had were with "companies" through job applications. So while a person may be responding to my fictitious job applications, it is really the companies that were the participants.

What safeguards are taken to relieve possible adverse consequences of these issues (e.g., informing participants about the study afterwards, extra safety regulations, etc.).

Participants were unaware that they were involved in an experiment. It was also impossible to attain informed consent, as informing participants would invalidate the study. However, direct evidence of labour market discrimination is not obtainable by any other method (Bertrand & Duflo, 2017; McGinnity et al., 2009). Additionally, List (2009) argues that when a study benefits society and ensures anonymity and just treatment, the absence of informed consent is justifiable.

Are there any unintended circumstances in the study that can cause harm or have negative (emotional) consequences to the participants? Indicate what possible circumstances this could be.

This methodology required the deception of fictitious candidates and their applications for jobs. To minimise the inconvenience for the employer or genuine applicants, I terminated the recruitment process when any job inquiries were made. Specifically, when an employer contacted both applicants, I withdrew the application as promptly as possible (within one day). However, if they only reached out to one applicant, I waited two days to remove the applications. McGinnity et al. (2009) contended that it is not likely that any employer would have endured substantial expenses or inconvenience with these procedures.

PART IV: SAMPLE

Where will you collect or obtain your data?

I responded to job openings within the Dutch labour market. I sent two resumes in response to each vacancy, both with either a post-secondary vocational education or vocational bachelor's degree, depending on the level of the job. I randomly assigned a photo of a black man and a white man to these resumes. In order not to raise suspicion, I applied with a difference of a couple of hours. The applications were to publicly advertised vacancies on online job markets (e.g., job portals).

The reaction of employers to the fictitious applicants was registered. Systematic differences in interview requests of the fictitious candidates provided the measure of racism (Riach & Rich, 2002). Responses were captured via specific email addresses or telephone inboxes for each applicant. All calls were immediately sent to voicemail, which was the standard message of the mobile provider.

What is the (anticipated) size of your sample?

I expected to respond to about 100 to 150 different vacancies. I responded to 150.

What is the size of the population from which you will sample?

The population includes all companies/job vacancies in the NL in business related fields. It is therefore not possible that any company will be singled out.

Part V: Data storage and backup

Where and when will you store your data in the short term, after acquisition?

On a USB stick with a password.

Who is responsible for the immediate day-to-day management, storage and backup of the data arising from your research?

Dennis van Doornik (myself)

How (frequently) will you back-up your research data for short-term data security?

Once a week

In case of collecting personal data how will you anonymise the data?

The outcome of this study was treated confidentially (Riach & Rich, 2004). No data was distributed relating to any individual business to ensure that the results would not directly impact any participants. So, organisation identifying information was not made available to any third party at any stage, including my thesis group and supervisor. In sum, I took all possible actions to minimise the individual risk and inconvenience to organisations and employers.